INTERPRETING RECENT RESEARCH ON SCHOOLING IN DEVELOPING COUNTRIES

Eric A. Hanushek

Policymakers in developing countries have long been troubled by the undesirable, but apparently unavoidable, choice between providing broad access to education and developing high-quality schools. Recent evidence, however, suggests that this is a bad way to think about human capital development. Grade repetition and high dropout rates lead to a significant waste of resources in many school systems. Students in quality schools, however, respond in ways that reduce such inefficiencies, perhaps even sufficiently to recoup immediately investments in quality.

Promoting high-quality schools, however, is more difficult than many have thought, in part because research demonstrates that the traditional approach to providing quality—simply providing more inputs—is frequently ineffective. Existing inefficiencies are likely to be alleviated only by the introduction of substantially stronger performance incentives in schools and by more extensive experimentation and evaluation of educational programs and school organizations. Incentives, decentralized decisionmaking, and evaluation are alien terms to education, in both industrial and developing countries, but they hold the key to improvement that has eluded policymakers pursuing traditional practices.

Recent research into schooling has begun to point consistently toward education policies that differ sharply from much of what we have seen in the past. In particular, it points more toward performance incentives and less toward regulatory and input-based policies, and it underscores the importance of developing high-quality schools, even if this goal appears to impinge on access to schools.

Three fundamental findings flow from the new research. First, education around the globe is a very inefficient exercise; strong evidence indicates that too much is being paid for the performance obtained from schools. Second, education has proved to be a very complicated subject, and available research has

The World Bank Research Observer, vol. 10, no. 2 (August 1995), pp. 227–46.

yielded little specific guidance on how to boost quality through standard regulatory and spending policies. This calls for fundamental changes in the way we conceptualize educational policy. In particular, various forms of performance incentives appear to offer more hope for improving schools and raising student achievement, even if we currently have little experience in developing such policies. Third, the importance of high-quality schools keeps reappearing in the research. Developing countries, in the interest of expanding the availability of schools, have tended to sacrifice quality. This approach, however, appears misguided, because students react to low-quality schools in ways that damage the very policy of expanded access. This paper examines these three findings and then outlines a set of natural policy measures that flow from these findings.

The Pervasiveness of Inefficiency

It is useful to begin with what is known about the effect of different resources on student performance. Traditionally much of school policy is an attempt at selecting an optimal set of resources—however defined—and ensuring that it is available. Matched with this policy perspective has been a line of research that reviews the relationship between resources and student performance. If the effectiveness of each resource were known, it would be straightforward to define an optimal set of resources and decide on the policies that would most likely produce high levels of educational achievement. Unfortunately, however, the information eludes us.

Although research into the determinants of students' achievement takes various approaches, one of the most appealing and useful has been what economists call the *production function* approach, which focuses on the relationship between school outcomes and measurable educational inputs. This research has been refined and used most extensively in the United States, but significant related evidence is becoming available about performance in developing countries.

The underlying model guiding the analysis of school performance is straightforward. It postulates that the output of the educational process—that is, the achievement of individual students—is related directly to a series of inputs. Policymakers directly control some of these inputs, such as the characteristics of schools, teachers, and curricula. But other inputs, such as the family's socioeconomic level and the student's innate endowment or learning capacity, are affected by public policy only indirectly, if at all. Further, although achievement is usually measured at discrete points in time, the educational process is cumulative; past inputs affect students' current levels of achievement.

Starting with this model, statistical techniques, typically some form of regression analysis, are used to identify the specific determinants of achievement and to make inferences about the relative importance of the various inputs to student performance. The accuracy of the analysis and the confidence warranted by the answers depend crucially on a variety of technical issues, related to mea-

surement and estimation, that need not be discussed here. Instead this summary highlights the overall findings of the model and shows how these results relate to potential schooling policies.

The Studies in This Analysis

Most such studies measure output on the basis of students' scores on standardized achievement tests, although a significant number of studies have used other quantitative measures, such as school attendance rates and school continuation or dropout rates. These quantitative measures are generally interpreted to be plausible indicators of future success in the labor market. (Several studies support this interpretation; see, for example, Behrman and Birdsall 1983; Jamison and Moock 1984; Boissiere, Knight, and Sabot 1985; and Knight and Sabot 1987, 1990.)

Empirical specifications have varied widely, but they have also had much in common. Family inputs tend to be measured by sociodemographic characteristics, such as parental education, income, wealth, and family size. Peer inputs, when included, are typically aggregate summaries of the sociodemographic characteristics of the other students. School inputs include the teachers' characteristics (level of education, experience, gender, religion, and so forth), the way schools are organized (class size, facilities, administrative expenditures), and community factors (such as average expenditures). Most empirical work in the United States has relied on data that were designed for other purposes, such as the schools' normal administrative records, but much of the work in developing countries is based on specialized data sets developed for studying schools.

Perhaps surprisingly, there is considerable similarity in the research findings across industrial and developing countries. These similarities are useful both in interpreting the results and in deriving policy implications. (For reviews of studies in industrial countries, see Hanushek 1986, 1989; Guthrie and others 1971; Averch 1974; Bridges, Judd, and Moock 1979; Murnane 1981; Glasman and Biniaminov 1981; and Murnane and Nelson 1984. For developing countries, see Fuller 1985; Fuller and Clarke 1994; Harbison and Hanushek 1992; and Velez, Schiefelbein, and Valenzuela 1993.) In all, close to 100 quantitative studies of the effectiveness of school resources in developing countries were available by 1991. Almost 400 studies for the United States were available by 1994.

The Findings

The central conclusion from a review of this literature is both simple and startling: in the last quarter century, work on the relationships between educational inputs and outputs has indicated that schools all over the world pursue very inefficient policies. On the question of efficiency, I rely on the simplest possible notion: do the resources purchased and used by the schools systematically improve student performance? At this stage I am concerned not with the

magnitude of the results, but only with whether the returns to the application of resources are positive. This statement about efficiency has strong implications for policy, but before delving into such implications, it is important to understand the nature of the evidence.

Table 1 summarizes the effects of five educational inputs on student performance in developing countries on the basis of ninety-six studies: teacher-pupil ratio; teacher education, experience, and salary; expenditure per pupil. A more recent review (Velez, Schiefelbein, and Valenzuela 1993) contains a larger number of studies, but the general conclusions are the same. Table 1 shows which inputs have a statistically significant correlation (by sign of coefficient or direction of effect) and which are statistically insignificant. (The insignificant findings, unfortunately, cannot be divided by direction of effect.) In all cases, the reported correlations are those that hold after allowing for differences in the family backgrounds of students and in other educational inputs.

The evidence provides no support for policies to reduce class size. Of the thirty studies investigating teacher-pupil ratios, only eight find statistically significant results supporting smaller classes; an equal number are significant but have the opposite sign; and almost half are statistically insignificant. These findings qualitatively duplicate those in the U.S. studies, but are particularly interesting here.[1] Class sizes in the developing-country studies are considerably more varied than those in the U.S. studies and thus pertain to a wider set of environments, providing even stronger evidence that the enthusiasm for policies to reduce class size is misplaced.

The effect of the teachers' experience yields results that are roughly similar to findings for the United States. Although 35 percent of the studies (sixteen out of forty-six) display significant positive benefits from more teaching experience (the analogous figure for the United States is 29 percent), the majority of the studies—twenty-eight out of forty-six—found this input statistically insignificant.

The results for teacher education, on the other hand, diverge in relative terms from those seen in the U.S. studies, with a majority (thirty-five out of sixty-

Table 1. *Summary of Ninety-Six Studies on the Estimated Effects of Resources on Education in Developing Countries*

Input	*Number of studies*	*Statistically significant*		*Statistically insignificant*
		Positive	*Negative*	
Teacher-pupil ratio	30	8	8	14
Teacher's education	63	35	2	26
Teacher's experience	46	16	2	28
Teacher's salary	13	4	2	7
Expenditure per pupil	12	6	0	6
Facilities	34	22	3	9

Source: Harbison and Hanushek 1992.

three) supporting the conventional wisdom that more education for teachers improves student performance. (In the U.S. studies, teachers' education was the least important of all inputs.) Although these results are still surrounded by considerable uncertainty (twenty-six estimates are insignificant and two display significantly negative effects), they do suggest a possible differentiation by stage of development and general level of resources available.

The evidence on teacher salaries in developing countries contains no compelling support for the notion that higher wages yield better teachers. Because these results aggregate studies across different countries, school organizations, and labor markets, however, it is difficult to take these results too far. For policy purposes, one would generally want information on what happens if the entire salary schedule is altered (as opposed to simply moving along a given schedule denominated, say, in experience, education, or some other attribute of teachers). But it is not possible with available studies to distinguish between the two effects.

Data on total expenditure per pupil are rarely available in analyses of developing countries, but the twelve studies that include such estimates are evenly split between statistically significant and statistically insignificant. Given questions about the quality of the underlying data, not too much should be inferred from these findings.[2]

One of the clearest divergences between the findings in developing and industrial countries is the effect of facilities, suggesting that differences in the school environment are of some importance in developing countries. Twenty-two of the thirty-four investigations support the provision of quality buildings and libraries. The specific measures of facilities vary widely, however, so the interpretation almost certainly depends on local conditions.

Several other factors have been investigated in the course of the developing-country analyses, including an assortment of curriculum issues, instructional methods, and teacher training programs. Many of these inputs, however, are difficult to assess here because of the multicountry evidence and the probable importance of local institutions. One input—the provision of textbooks—has received widespread endorsement, although this support is as much for conceptual reasons as for solid empirical ones. The relationship of textbooks and writing materials to student performance is found to be important with reasonable consistency in developing countries, but relatively few studies are available (Lockheed and Hanushek 1988; Lockheed and Verspoor 1991). Investigations of technological or organizational differences have shown mixed results. In three extensive investigations in Nicaragua, Kenya, and Thailand, interactive radio teaching, an approach to "distance education," has been found to be effective in teaching children in sparse settlements in rural areas. This result should not be generalized to all new technologies, however. In particular, there is little evidence at this time to support the widespread introduction of computers (Lockheed and Verspoor 1991).

Although available studies cover a wide range of circumstances and inputs into schooling, the standards of data collection and analysis are so variable that

the results from this work are subject to considerable uncertainty. Much of the analysis of input-output relationships for developing countries is not published in standard academic journals, and thus it lacks that basic level of quality control. Even more important, the data for many of these studies do not come from regular collection schemes, are difficult to check for quality, and miss key elements of the educational process. Therefore, even if the analytical approaches are state of the art, many questions remain.

To supplement these findings, I have noted the similarity of the work in developing countries to that done in the United States, where search produces almost 400 separate studies that relate resources to student performance. Harbison and Hanushek (1992); Hanushek, Rivkin, and Taylor (1995); and Hanushek (forthcoming) come to many of the same conclusions. Little evidence suggests that smaller classes are better than large classes. Additionally, things that affect teacher salaries—particularly education and experience—are not systematically related to student performance. Neither are teacher salaries or aggregate differences in per-pupil expenditures. Schools in both developing and industrial countries continue to pay for things that have little consistent or systematic payoff in terms of student performance. In simplest terms, policies that are expensive but that have no return are obviously inefficient.

These findings indicate that there are not clear and systematic relationships between key inputs and student performance. This does not say that there never is such a relationship. In fact there is reason to believe that each of these resources is sometimes productive. The data in table 1 demonstrate that the estimated effect of each of the resources is positive and statistically significant more frequently than one would expect if the underlying relationship were *always* zero; that is, that the frequency of positive and statistically significant estimates is greater than the 5 percent that would be expected from random occurrence when the true relationship is identically zero. Formal statistical analyses confirm that resources have been used effectively. For U.S. studies, Hedges, Laine, and Greenwald (1994) have received considerable attention. An approximate version of their analysis is presented by Michael Kremer in his response to this article. Even though there are important concerns about both of these statistical manipulations (because of publication bias and dependence of the estimated coefficients), there is no disagreement about the basic point: resources are sometimes employed effectively. But quite frequently these resources are ineffective. Both Hedges, Laine, and Greenwald and Kremer, while concentrating on the evidence against *complete* lack of effects, clearly state that what matters most is how money is spent (see Hanushek 1994 for a discussion of statistical and policy issues). And, importantly, there is no simple description of when these resources will be effective.

This analysis does not say that differences in resources could never be important, just that they have not been, given the way schools are organized. Part of this could, in the case of developing countries, reflect variations across countries. Certain countries may organize their school systems to promote perfor-

mance and efficient resource usage. These effective countries might be lost in the pool of all countries where others are quite ineffective. Here, however, the similarities of results with the United States appear instructive, because the U.S. variation does not appear to reflect any simple differences in school organization.

The analysis by Card and Krueger (1992) of the effects of historical resource usage in the United States, with its observations of schools during the 1920s and 1930s, may capture some differences in organization or level of resources that would be more appropriate for developing countries. Intense controversy surrounds the interpretation of their results, however, making it difficult to apply the work to investment policies in developing countries. (For a discussion of Card and Krueger 1992, see Betts 1994, forthcoming; Heckman, Layne-Farrar, and Todd 1994; Hanushek, Rivkin, and Taylor 1995; and Speakman and Welch 1995).

These findings do not imply that all schools are the same. Quite the contrary, schools differ dramatically. These results suggest that the measurable factors, factors that often determine central policy, are not consistently related—a subject that is discussed in the next section.

Some evidence suggests that minimal levels of basic school resources, such as the availability of textbooks, the provision of minimal facilities, and so on, are important in student achievement (Lockheed and Verspoor 1991; Lockheed and Hanushek 1988; Harbison and Hanushek 1992). These findings are not uniform in the statistical analyses, but they are common enough to receive more attention than most of the other findings. Why then do we not see policies instituted to provide these minimal resources, particularly if there is reasonably strong evidence about their importance?

The Complexity of Education

If providing more resources will not reliably lead to improvements in student performance, what alternatives do policymakers have? The preceding findings have led to several continuing strands of research designed to aid the development of better policies. Unfortunately, these efforts have not provided the information needed to refine traditional input-centered policies. Nor is the extension of existing studies likely to improve our knowledge in a timely and satisfactory way.

One line of research has attempted to add other measures of how schools work—the educational process, the preparation of teachers, and various other measurable factors related to schools and education. A recent, and particularly rich, example of research that reviews a wide range of inputs (Glewwe and others 1995) suggests that student achievement responds to various pedagogical factors, such as the amount of time devoted to instruction or the use of written assignments. By looking at both "resource" and "process" factors, the authors

suggest that a broader view is preferable to the more narrow concentration on resources that has typically been pursued. (This work corresponds to a recent review of school performance in developing countries by Fuller and Clarke 1994.)

The idea behind this approach is that previous research has not yet found the right descriptors of education and schools. The overall notion is that we can build a more complete picture by adding measures that appear to represent important differences and then use that information to develop more refined policies. I am not, however, persuaded about the usefulness of either the research strategy or the policy implications that would necessarily follow.

Almost all studies that include enriched descriptions of schools, teachers, organization, and pedagogy find that a set of these factors is significantly related to student performance. These studies note that richer characterization of schools better explains student achievement than resources alone and conclude that policies directed at the factors uncovered (in each individual study) could be beneficial. There are different ways to look at these data, however.

First, simply because these expanded studies appear to explain student performance better than measured resources alone, they still might not take us very far. We have fairly conclusive evidence that measured resources are not systematically related to student performance. Thus the correct comparison is not between pedagogical and organizational measures and resource measures, but between these and differences in teacher quality and school quality, whether measured or not. This distinction is very important because the measured resources do not give a clear indication of the underlying differences among schools. Do additional measures of schools explain a large portion of the systematic differences across schools? None of the available studies of other school factors address this issue, in part because they do not identify the overall systematic variations in school quality.

Second, many individual studies find "strong" evidence that some specific factor, say, teacher training or the structure of instructional time, is important to student achievement. This kind of finding is readily seen in table 1. For example, thirty-five studies (out of sixty-three) find positive and statistically significant relationships between teacher education and student performance. Each of these studies, taken by itself, would give "strong" evidence. But looked at from the opposite vantage point, an almost equal number—twenty-eight—do not support the conventional wisdom. This example suggests that some skepticism is required in looking at any individual study, no matter how good the study is in a scientific sense. Other factors in table 1 are similarly ambiguous. The importance of the inherent uncertainty can frequently be readily seen by looking at the separate estimated educational production functions often contained within a given analysis. For example, some analyses contain separate estimates of how specific resources affect different measures of achievement, say math and reading, for a given sample of students. Within such analyses, where presumably the separate estimated relationships are highly correlated (because they involve the same students, the same schools, and so forth), there is

frequently little confirmation of any specific findings for resources or measured pedagogical factors. With one measure of outcome, for example, the education of the teacher may be important, while the same might not be true for another measure of outcome.

Finally, generalizations are difficult because the factors that are considered are quite idiosyncratic. Individual studies, because of peculiarities in the available data and the varying perspectives of the researchers, tend to pursue different measures of school and teacher factors. Each generally highlights its "new findings." It would, of course, be foolhardy to think of making policy on the basis of these individual factors (and few of the original authors would suggest doing so).

My own interpretation of existing evidence, based on results for both the United States and developing countries, is that schools differ in important ways, but we cannot describe what causes these differences very well. To take one example, Hanushek and Lavy (1994) investigated differences in the quality of schools across a sample of primary schools in Egypt. We defined school quality implicitly. After allowing for individual differences among students in achievement and in parental education, we labeled schools that had large gains in student achievement in a given year as high-quality schools; those with small gains, low quality. A continuous measure of school quality was developed by looking at growth in student achievement (after considering family and other influences on achievement growth). This exercise found enormous differences in the sixty sample schools. Table 2 shows the variation in the quality of schools by looking at achievement relative to a randomly chosen base school. The worst school shows an average achievement gain that is 62 percent below the base school, while the best school is 30 percent above. These results indicate dramatically that schools do differ in quality and that the difference is enough to be relevant to policymaking.

At the same time, measured attributes of teachers and schools explain only a small portion of these differences. From our estimation, only 16 percent of the variance in school quality is related to teacher attributes (such as education and gender) and school attributes (such as class size and facilities). Although we did not look further, I seriously doubt that adding more detailed measures of resources, or of pedagogy, or of curricular differences would have allowed us to explain the differences much more fully.

Table 2. *Distribution of Estimated School Quality in Egyptian Primary Schools*

Distribution	*All schools*	*Rural*	*Urban*
Mean	−.084	−.111	−.057
Minimum	−.62	−.62	−.52
Maximum	.30	.30	.21

Note: Values indicate the average proportional achievement gain of a school in comparison with that of the arbitrarily chosen base school, Taha Hussein School.

Source: Hanushek and Lavy 1994.

A similar approach undertaken in rural Brazil (Harbison and Hanushek 1992) pointed to very similar conclusions: schools show very large differences in their ability to improve student achievement, but these differences are not highly correlated with measured characteristics of teachers and schools.

In short, the findings summarized in table 1 do *not* indicate that schools and teachers are all the same. Large differences exist, even though these differences are not captured by the simple measures commonly employed. Neither, it appears, are they captured by more detailed measures of classroom organization or pedagogical approach. This leads me to conclude that the educational process is very complicated and that we do not understand it very well. We cannot describe what makes a good or bad teacher or a good or bad school. Nor are we likely to be able to describe the educational process very well in the near future. My view is that we should learn to live with that fact: living with it implies finding policies that acknowledge and work within this fundamental ignorance.

Quality versus Access

A third major aspect of current research relates to the importance of school quality and particularly to the perceived policy tradeoff between quality and access. The traditional concern goes something like this: given limited budgets for schools, and the commonly accepted twin objectives of expanding access and improving quality, policymakers face a particularly unpleasant dilemma. They must choose between expanding the availability of education or providing high-quality schools.

A second way of viewing these policy concerns, while apparently different, is actually quite closely related. Analyses of labor market implications and the rate of return to schooling in developing countries suggest strongly that schooling is a very good investment. A year of schooling typically shows a 25–30 percent real rate of return, which appears noticeably better than that of other investment alternatives. At the same time, school completion rates in low-income countries are very low (Lockheed and Verspoor 1991). These two facts are inconsistent. If education yields such a high rate of return, why are people not taking advantage of it?

Emerging analyses of school quality have something to say about both elements of education policy. I believe that the common conception of a simple tradeoff between access and quality is misleading—if not wrong; and I think that low school quality may frequently be an important explanation for the widespread failure to take advantage of the apparently high returns available from education.

The central theme here is that school quality is directly related to students' decisions about attending school and schools' decisions about promoting students. High-quality schools raise student achievement and speed students through primary (and perhaps secondary) school, thus saving costs. Additionally, stu-

dents respond to higher school quality with lower dropout rates: they tend to stay in good schools and drop out of poor ones.

Both of these mechanisms indicate a direct relationship between the quantity of schooling attained and the quality of that schooling. Thus, studies of the rate of return to schooling that consider only the quantity of schooling produce a misleading estimate of the potential gains. Estimates of the rate of return to schooling that do not account for quality differences will systematically overstate the productivity gains that are associated with additional years of schooling, because the estimates will include quality differences that are correlated with quantity. The evidence shows that those who do not complete a given level tend to have attended poorer schools. If a policy simply pushes students to stay in school but makes no changes in the fundamental quality of the schools, the new school completers will get only the returns associated with years of schooling and not with quality. Thus, their rate of return on their investment in schooling will not be as high as the estimates suggest.

Many countries, concerned about very high grade repetition rates, directly intervene to ensure regular promotion through school (Lockheed and Verspoor 1991), but they typically ignore school quality. Neglecting the quality of schools is a serious mistake. In studying primary school students in the rural northeast of Brazil, Ralph Harbison and I discovered a very direct relationship between what a student knows and the student's promotion probabilities. Students who learn more than the curriculum requires (as measured by specifically designed tests) are significantly more likely to be promoted through primary school than those who do not learn what is expected. Schools, not surprisingly, have an important impact on student achievement. These findings suggest that policies that improve the quality of schools—that is, that enhance student achievement—will simultaneously lead to more rapid progress by students through the grades.

The magnitude of the overall effects of improving school quality, when converted to a monetary metric, is remarkable. Hanushek, Gomes-Neto, and Harbison (1994) summarize the expected savings from two simple policies—improving the availability of textbooks and writing materials (software) or improving components of the facilities (hardware). They show that if $1 is invested in useful resources such as textbooks, an immediate savings of more than $12 is obtained from speeding students through school. (These savings are pure efficiency savings from getting through school more quickly and include none of the increased productivity benefits that typically justify schooling investments; increases in future productivity simply reinforce the efficiency gains.) Where facilities are lacking, each $1 improvement has an expected cost saving of more than $3.

These estimates of the savings that can be expected from quality improvements are subject to some uncertainty. Nonetheless, the lowest plausible savings still indicate substantial efficiency gains from improving the quality of schools. The availability of books and writing materials and school facilities is consistently important for student achievement and promotion.

These results highlight the importance of providing minimal resources for schools and are consistent with previous findings about the importance of basic textbooks, materials, and facilities. But these estimates—as startling as they are—may not represent the largest opportunities and are really lower-bound estimates of the potential for change. Specifically, all the research points to the importance of the teacher.[3] Because the variations in teacher quality appear to be much more important than the variations in software or hardware, the savings from ensuring the former would almost certainly exceed those obtained from improvements in the latter. Unfortunately, because we do not know how to hire particularly effective teachers—nor what it would cost—we cannot calculate straightforward benefit-cost ratios.

Grade repetition is not entirely bad, because students do learn more with each time through the same grade, but it is an expensive way to improve student learning (Gomes-Neto and Hanushek 1994). One alternative explanation is that repetition reflects demand-side factors—that is, student choices that lead to low attendance during each school year. Little is actually known about attendance patterns, but anecdotal evidence suggests that normal crop cycles and requirements for children to work in the fields at planting and harvest times may be important in some settings. Such attendance patterns could severely constrain the chances of completing a given grade, at least in the likely absence of well-integrated, self-paced instruction. Dealing with these issues might require different policies aimed at lessening the current consumption constraints of families.[4] In any event, however, the continued production of low-grade schools is no more effective in the face of such demand-side influences than without them.

In work on Egypt, Hanushek and Lavy (1994) pursue a related question: whether school quality affects students' decisions to drop out. In that analysis, the school quality estimates (table 2) were included as one of the determinants of the decisions of individual students. Additionally, the analysis considered the students' own achievements and abilities as well as their earnings opportunities outside of school. If we hold achievement and opportunities constant, students going to high-quality schools are much more likely to stay in school than those going to low-quality schools. This makes sense. If a student is not going to get anything out of school, why waste the time?

The magnitude of the effect is particularly important. The primary schools sampled had average dropout rates in 1980 of 9.3 percent. If all the schools were at the quality level of the best one, the dropout rate would fall to 3.2 percent or less, a decline that indicates the huge impact of quality on school attainment.

Research in Brazil and Egypt points to similar conclusions. School quality has large and direct effects on school access and school attainment. These effects are complements, not substitutes, as suggested by the simple budgetary analysis that is commonly employed. And the research in both countries indicates that quality adds a dimension that is extremely important in thinking about schooling in developing countries. Finally, efforts to pursue quality improvement must

confront the policy challenges described in the earlier sections. Inefficiency and general lack of knowledge about the production function in education imply that dealing with quality will require new and innovative approaches.

These conclusions are supported in Glewwe and Jacoby (1994), whose work on Ghana shows the direct relationship between school quality and school attainment. Improving the schools (in this case, the facilities) tends to hold students in school longer, other things equal. The authors do not obtain estimates of the total effects of school quality (as was done for Egypt), but the indication that measured effects have this influence confirms the quantity-quality correlation. This correlation in turn confirms the bias in rates of return flowing from analyses that ignore variations in school quality.

A Policy Perspective

From a policy perspective, the central reason for pursuing research on educational performance has been to develop a list of inputs, curricular elements, or other factors affecting student achievement that could be instituted through central policy. If there were a clear understanding of what determined student performance, individual schools could be told what to do. Indeed, central authorities could insist that schools follow their directives—by linking funding to specific actions, regulating certain approaches, or implementing specific hiring decisions. But we do not understand what determines educational achievement, and, in my view, the research indicates that such prescriptive policies are the wrong approach.

If pursuing such policies is unlikely to lead to overall improvements, what can be done? Several policies appear to be viable alternatives if we shift our approach from policies based on input to those based on performance incentives. Performance-based policies are those that reward accomplishment—such as good reading skills or adequate numeracy skills. These policies would specify end goals, provide carrots and sticks related to them, and harness the energies of the actors in the system, but they would not specify how individual schools should achieve these goals.

What kinds of incentive systems fit into this category? Merit pay that rewards teachers for what students learn is a simple example; it is also a system that has recorded little historical success (see Cohen and Murnane 1986). Other variants of performance-based rewards within current systems have been suggested, although rarely implemented or evaluated. More radical changes could also provide improved incentives, although we have little experience with them. Under private contracting arrangements, for example, contracts are awarded to people outside the school system who are hired to run certain parts of the schools and are rewarded on the basis of their effectiveness in achieving the stated objectives. Offering parents a choice of schools or providing a voucher system are more extreme versions of this mechanism. Parents choose schools, and the con-

comitant shift in the flow of resources provides the incentive to produce what the parents desire. The unifying theme is that good performance is rewarded while poor performance is not.

This message is not aimed solely at developing countries; industrial countries are also beginning to respond to it. Hanushek and others (1994) described the recommendations of a panel of economists on introducing economic logic into the reform of U.S. schools.

This is also not a particularly pleasing message for policymakers because it requires considerable change. Policymakers frequently wish to maintain control of policies and distrust the motives and abilities of local decisionmakers. Thus, they frequently look for explicit, centralized approaches. Accepting the decentralized decisionmaking that almost certainly accompanies performance incentives requires a very different focus.

Moving to performance incentives involves considerable uncertainty. We have substantial evidence that the current structure is not working effectively, but we have little experience with the alternatives. Additionally, to apply such a system, we must be able to specify desirable goals and to measure performance toward those goals—a difficult, sometimes contentious, process that will take considerable experimentation and evaluation. These are not procedures most schools, in either industrial or developing countries, do routinely or easily.

A variety of nascent experiments, such as the use of vouchers in Chile, Colombia, and the United States (Milwaukee, Wisconsin), are being conducted. Given current knowledge, however, we do not really know much about instituting such approaches. I would therefore recommend a plan of systematic experimentation and evaluation with alternatives. The full description of alternatives, of experimentation, and of evaluation approaches goes considerably beyond the scope of this paper (see Hanushek with others, 1994). More energy must go into developing new organizational forms if education in developing countries is to improve.

It is unlikely that we can accomplish such improvement through current organizations and through regulatory structures that dictate specific remedies. For example, textbooks. Research in past decades has shown that improved textbook availability tends to increase student performance, and yet current incentives have not uniformly directed additional expenditures toward the purchase of books and writing materials. Such funds may be spent on learning materials, or they may be spent elsewhere. Several explanations come to mind. The findings may be inaccurate, confusing textbooks with other characteristics of schools, parents, and communities. The findings may be correct, but policymakers may have other objectives than efficiently increasing student achievement. Or the findings may be correct, but other deficiencies are more important than textbook shortages. Depending on which explanation is correct, mandating greater textbook availability may or may not lead to improved performance. On the other hand, providing strong incentives with information from the research may be expected to lead local schools to make appropriate decisions about textbooks

or other spending. Although the latter presumption is uncertain because of our lack of experience with performance incentives, incentives do offer a more likely possibility for improvement.

Evaluation and Innovation

One consistent observation about schools around the world is that the education system itself learns very slowly. Although educators are dedicated to teaching students, they are reluctant to submit to the often-painful process of evaluation and learning. Therefore, new ideas are seldom subjected to thorough evaluation, nor are active decisions often made about their success or failure. Because of the need to expand education and improve school systems in developing countries, many informal experiments are currently under way. Indeed, the World Bank and other international agencies that support the improvement of education—both financially and intellectually—frequently insist on new approaches to the organization and conduct of schooling. Yet little systematic information is collected about these efforts, and even less of an attempt is made to evaluate the information that is collected.

This state of affairs might be understandable and acceptable if student achievement were better. If international organizations and individual countries tended to make good decisions about the effective use of resources and the provision of education, there would be far less concern about evaluation. But the evidence suggests that the process is not one that leads to confidence. The programs of the World Bank, which have increasingly been focused on the human capital needs of developing countries, provide a natural array of interventions to education that could, conceptually, provide sorely needed information about productive lines of improvement. The Bank regularly requires the evaluation of loans in which it participates, but unfortunately these evaluations seldom involve any detailed analytical work that would permit dissemination of new techniques or new organizational forms.

Much more serious assessment efforts will be required. Many countries do little systematic testing and evaluation of student performance. Clearly this is not a small issue because developing and evaluating appropriate testing and measurement methods are themselves the subject of intense discussion. This step cannot be ignored if nations are to develop the knowledge and organization required for effective educational systems.

Implementation Issues

The simple explanation for the current state of inefficiency is that few if any incentives are aimed at improving student performance by adding resources. When there are no direct incentives to increase achievement and when education is so heavily influenced by actions in the individual classroom and school, it

is not particularly surprising that the system responds as it does. Added resources may or may not be converted into improved performance depending on specific local personnel and factors.

But this explanation begs the larger questions about why school systems do not have built-in performance incentives. Although most school systems are publicly financed, public financing alone would not seem to preclude the development of incentive systems. The common explanations center on some combination of bureaucratic incentives and union restrictions. The assumption is that these actors are pursuing their own self-interest or are motivated by the desire to control decisionmaking authority. Little explicit analysis exists to support or refute this assumption, however.

Understanding the source of inefficiency is important because improvement requires changes in the existing structure. If the inherent source of inefficiency and of resistance to change is the self-interest of current actors, any reform program must take this into account. For instance, current teachers must be incorporated into any program of change. The specific approach depends on the new organizational structure that is introduced, but an important element might be protecting existing teachers from arbitrary changes. One way of moving to new, more responsive systems would be to offer two-tier contracts in which current teachers retain their contracts and new teachers are subject to contracts with very different incentives.

Fair and equitable treatment of existing school personnel is essential even if the long-run plan is to replace the majority of existing employees. Existing teachers will be important in the transition because adjustments will be built upon the existing system. And, given the continued importance of the individual teacher and headmaster, forcing the adoption of new methods from the top down is typically not possible without also introducing considerable inefficiency.

If the root problem involves existing government bureaucracy, it is more difficult to see how to proceed because structuring appropriate incentives for bureaucracies appears even more difficult than devising incentives for teachers. Strong leadership from the top is frequently an essential element of reform. The efficacy of providing incentives for public officials depends on the specific circumstances of individual countries and local structures, but obviously nothing is likely on a widespread basis unless the key governmental actors set performance goals. There is, of course, a qualitative difference between teachers and bureaucracies because teachers have much more direct control over student performance and thus are more directly affected by student performance incentives. The bureaucracy has a more indirect role through facilitating the performance of schools and monitoring the actions of local schools. Thus, providing leadership incentives requires complicated reward structures. Government agencies seldom offer much in the way of performance rewards, partly because they pay little attention to measuring outcomes. The tendency toward insulating government agencies from performance incentives is frequently reinforced by public employee unions that resist the individual differentiation of workers.

Could the introduction of competition initiate implicit incentives to reform school policies? The conceptual basis for educational vouchers, which permit students to go to either public or private schools, is that competition for students (and the associated resources for their education) could force schools to be more responsive to student achievement and more efficient. At the same time, educational vouchers might also offer competition to much of the bureaucracy involved in directly providing public schooling, which could have similar incentive effects for bureaucrats and teachers. (The bureaucracy involved with administering and monitoring the vouchers would not feel these competitive pressures.)

These issues of political economy and the existing school personnel cannot be ignored if the organizational and incentive structures in the schools are to be changed. It is unrealistic to assume that the staff will automatically adopt new policies that increase the risks of unemployment, that eliminate some of the implicit benefits in their current employment, and that potentially require new techniques and more work. Either these people must be brought into the process of change or mechanisms must be developed to ensure that they do not sabotage any changes.

As noted earlier, most incentive systems require mobilizing local initiative. In a wide variety of situations today, local schools do not appear to have sufficient capacity to invent and operate new organizations with altered incentives. Little is known about how best to develop such capacity, but clearly trying to initiate change through central directives is quite difficult.

Conclusions

The research into the educational process, both in the United States and in the developing world, promises some very distinct payoffs. In policy dimensions, we appear to have learned a great deal. At the same time, the results do not always conform to what was expected. Research conclusively demonstrates an inefficiency in the current organization of schools. Resources are being spent in unproductive ways—ways that do not contribute to improving student performance. Correcting these inefficiencies is not simple. There is no blueprint for a model school that can be reproduced and handed out to policymakers, and such a blueprint is unlikely to be developed in the near future. Instead, we must turn to new organizations and new incentives if we are to improve schools.

Research suggests that the most likely path to improvement involves the introduction of performance incentives. Although several ways to introduce incentives have been suggested, none has been tried extensively. An extensive and systematic program of experimentation and evaluation is thus in order.

Finally, the evidence underscores the importance of establishing good schools. Although translating this goal into policy will be difficult, there are powerful reasons to believe that providing quality schools should be very high on the

policy agenda. The continued expansion of low-quality schools—often thought to be a step on the path both to high access and to high-quality schools—may actually be a self-defeating strategy.

Notes

Eric A. Hanushek is professor of economics and public policy and Director of the Wallis Institute of Political Economy at the University of Rochester. This is a revised version of a paper presented in Washington, D.C., at the World Bank Conference on Human Capital and Development in May 1994. It has benefited from helpful suggestions from Harold Alderman, Victor Lavy, and anonymous referees.

1. The U.S. evidence about the unimportance of class size is even stronger. In 277 separate estimates of the effects of teacher-pupil ratios, only 15 percent have a positive and statistically significant impact on student performance (Hanushek, Rivkin, and Taylor 1995; Hanushek, forthcoming). A similar proportion shows negative and statistically significant impacts. The econometric evidence is consistent with experimental evidence about the lack of a consistent relationship between class size and student performance (Glass and Smith 1979, Word and others 1990). Some debate about school resources and subsequent earnings follows from Card and Krueger 1992, who link teacher-pupil ratios and other school inputs to wages. Recent discussions, however, raise serious questions about the estimated relationships found there (Betts 1994; Heckman, Layne-Farrar, and Todd 1994, Speakman and Welch 1995).

2. Although expenditure data are more plentiful in the U.S. studies, the quality of these analyses is low (Hanushek, Rivkin, and Taylor 1995).

3. The basic findings of variations across teachers come from attempts to estimate "total" teacher effects through general covariance models. In these models, variations in the average growth of student achievement across teachers is compared. The most conclusive estimates of variations in teacher quality are found in studies of U.S. schools (for example, Hanushek 1971, 1992; Murnane 1975; and Armor and others 1976). The analysis of Brazil considers differences across schools, but, because the schools are small, these differences are frequently associated with individual teachers (Harbison and Hanushek 1992).

4. Regular absence from school means forgoing high rates of return to human capital production. This behavior is presumably the outcome of borrowing constraints on households and current consumption needs. An alternative policy would involve loans that are contingent upon the attendance and successful completion of schooling. Because schooling has a high rate of return, such loans could be repaid through higher future earnings, thus aligning repayment with returns. To my knowledge, however, no such programs exist in developing countries.

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Armor, David, Patricia Conry-Oseguera, Millicent Cox, Niceima King, Lorraine McDonnell, Anthony Pascal, Edward Pauly, and Gail Zellman. 1976. *Analysis of the School Preferred Reading Program in Selected Los Angeles Minority Schools.* Santa Monica, Calif.: Rand Corp.

Averch, Harvey A. 1974. *How Effective is Schooling? A Critical Review and Synthesis of Research Findings.* Englewood Cliffs, N.J.: Educational Technology Publication.

Behrman, Jere R., and Nancy Birdsall. 1983. "The Quality of Schooling: Quantity Alone is Misleading." *American Economic Review* 73(5):928–46.

Betts, Julian R. 1994. "Is There a Link Between School Inputs and Earnings? Fresh Scrutiny of an Old Literature." University of California, Department of Economics, San Diego. Processed.

_____. Forthcoming. "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth." *Review of Economics and Statistics.*

Boissiere, Maurice X., John B. Knight, and Richard H. Sabot. 1985. "Earnings, Schooling, Ability, and Cognitive Skills." *American Economic Review* 75(5):1016–30.

Bridges, R. Gary, Charles M. Judd, and Peter R. Moock. 1979. *The Determinants of Educational Outcomes: The Impact of Family, Peers, Teachers and Schools.* Cambridge, Mass.: Ballinger.

Card, David, and Alan B. Krueger. 1992. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy* 100(1):1–40.

Cohen, David K., and Richard J. Murnane. 1986. "Merit Pay and the Evaluation Problem: Understanding Why Most Merit Pay Plans Fail and a Few Survive." *Harvard Educational Review* 56(1):1–17.

Fuller, Bruce. 1985. *Raising Quality in Developing Countries: What Investments Boost Learning?* Washington, D.C.: World Bank.

Fuller, Bruce, and Prema Clarke. 1994. "Raising School Effects While Ignoring Culture? Local Conditions and the Influence of Classroom Tools, Rules, and Pedagogy." *Review of Educational Research* 64(1):119–57.

Glasman, Naftaly S., and Israel Biniaminov. 1981. "Input-Output Analyses in Schools." *Review of Educational Research* 51(4):509–39.

Glass, Gene V., and Mary Lee Smith. 1979. "Meta-analysis of Research on Class Size and Achievement." *Educational Evaluation and Policy Analysis* 1(1):2–16.

Glewwe, Paul, Margaret Grosh, Hanan Jacoby, and Marlaine Lockheed. 1995. "An Eclectic Approach to Estimating the Determinants of Achievement in Jamaican Primary Education." *The World Bank Economic Review* 9(2, May):231–58.

Glewwe, Paul, and Hanan Jacoby. 1994. "Student Achievement and Schooling Choice in Low Income Countries: Evidence from Ghana." *Journal of Human Resources* 29(3):841–64.

Gomes-Neto, João Batista, and Eric A. Hanushek. 1994. "Causes and Consequences of Grade Repetition: Evidence from Brazil." *Economic Development and Cultural Change* 43(1):117–48.

Guthrie, James W., George B. Kleindorfer, Henry M. Levin, and Robert T. Stout. 1971. *Schools and Inequality.* Cambridge, Mass.: MIT Press.

Hanushek, Eric A. 1971. "Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data." *American Economic Review* 60(2):280–88.

_____. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* 24(3):1141–77.

_____. 1989. "The Impact of Differential Expenditures on School Performance." *Educational Researcher* 18(4):45–51.

_____. 1992. "The Trade-off Between Child Quantity and Quality." *Journal of Political Economy* 100(1):84–117.

_____. 1994. "Money Might Matter Somewhere: A Response to Hedges, Laine, and Greenwald." *Educational Researcher* 23(4):5–8.

_____. Forthcoming. "The Quest for Equalized Mediocrity: School Finance Reform without Consideration of School Performance." In Lawrence O. Picus, ed., *Where Does Money Go? Education Reform and the Allocation of Resources in Public Schools.* Newbury Park, Calif.: Corwin Press.

Hanushek, Eric A., with Charles S. Benson, Richard B. Freeman, Dean T. Jamison, Henry M. Levin, Rebecca A. Maynard, Richard J. Murnane, Steven G. Rivkin, Richard H. Sabot, Lewis C. Solmon, Anita A. Summers, Finis Welch, and Barbara L. Wolfe. 1994. *Making Schools Work: Improving Performance and Controlling Costs.* Washington, D.C.: The Brookings Institution.

Hanushek, Eric A., João Batista Gomes-Neto, and Ralph W. Harbison. 1994. "Self-financing Educational Investments: The Quality Imperative in Developing Countries." University of Rochester, Department of Economics, Rochester, N.Y. Processed.

Hanushek, Eric A., and Victor Lavy. 1994. *School Quality, Achievement Bias, and Dropout Behavior in Egypt.* Living Standards Measurement Study Working Paper 107. Washington, D.C.: World Bank.

Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor. 1995. "Aggregation Bias and the Estimated Effects of School Resources." Working paper. University of Rochester, Rochester Center for Economic Research, Rochester, N.Y. Processed.

Harbison, Ralph W., and Eric A. Hanushek. 1992. *Educational Performance of the Poor: Lessons from Rural Northeast Brazil.* New York: Oxford University Press.

Heckman, James J., Anne S. Layne-Farrar, and Petra E. Todd. 1994. "Does Measured School Quality Really Matter?" University of Chicago, Department of Economics. Processed.

Hedges, Larry V., Richard D. Laine, and Rob Greenwald. 1994. "Does Money Matter? A Metaanalysis of Studies of the Effects of Differential School Inputs on Student Outcomes." *Educational Researcher* 23(3):5–14.

Jamison, Dean T., and Peter R. Moock. 1984. "Farm Education and Farm Efficiency in Nepal: The Role of Schooling, Extension Services and Cognitive Skills." *World Development* 12(January):67–86.

Knight, John B., and Richard H. Sabot. 1987. "Educational Policy and Labour Productivity: An Output Accounting Exercise." *The Economic Journal* 97:199–214.

_____. 1990. *Education, Productivity, and Inequality.* New York: Oxford University Press.

Lockheed, Marlaine E., and Eric A. Hanushek. 1988. "Improving Educational Efficiency in Developing Countries: What Do We Know?" *Compare* 18(1):21–38.

Lockheed, Marlaine E., and Adriaan Verspoor. 1991. *Improving Primary Education in Developing Countries.* New York: Oxford University Press.

Murnane, Richard J. 1975. *Impact of School Resources on the Learning of Inner City Children.* Cambridge, Mass.: Ballinger.

_____. 1981. "Interpreting the Evidence on School Effectiveness." *Teachers College Record* 83(1):19–35.

Murnane, Richard J., and Richard R. Nelson. 1984. "Production and Innovation When Techniques are Tacit: The Case of Education." *Journal of Economic Behavior and Organization* 5(September/December):353–73.

Speakman, Robert, and Finis Welch. 1995. "Does School Quality Matter? A Reassessment." Texas A & M University, College Station, Tex.

Velez, Eduardo, Ernesto Schiefelbein, and Jorge Valenzuela. 1993. "Factors Affecting Achievement in Primary Education." HROWP Working Paper 2. World Bank, Department of Human Resources Development and Operations Policy, Washington, D.C. Processed.

Word, Elizabeth, John Johnston, Helen Pate Bain, B. DeWayne Fulton, Jayne Boyd Zaharies, Martha Nannette Lintz, Charles M. Achilles, John Folger, and Carolyn Breda. 1990. *Student/Teacher Achievement Ratio (STAR), Tennessee's K-3 Class Size Study: Final Summary Report, 1985–1990.* Nashville: Tennessee State Department of Education.